



Published in final edited form as:

*Biol Psychiatry*. 2022 February 15; 91(4): 389–398. doi:10.1016/j.biopsych.2021.01.015.

## Brain DNA Methylation Patterns in *CLDN5* Associated with Cognitive Decline

Anke Hüls, PhD<sup>1,2</sup>, Chloe Robins, PhD<sup>3</sup>, Karen N. Conneely, PhD<sup>1</sup>, Rachel Edgar, BSc<sup>4</sup>, Philip L. De Jager, MD, PhD<sup>5,6</sup>, David A. Bennett, MD<sup>7</sup>, Aliza P. Wingo, MD<sup>8,9,#</sup>, Michael P. Epstein, PhD<sup>1,#</sup>, Thomas S. Wingo, MD<sup>1,3,#</sup>

<sup>1</sup>Department of Human Genetics, Emory University, Atlanta, Georgia, USA

<sup>2</sup>Department of Epidemiology and Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

<sup>3</sup>Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

<sup>4</sup>Centre for Molecular Medicine and Therapeutics, BC Children's Hospital Research Institute, and Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada

<sup>5</sup>Cell Circuits Program, Broad Institute, Cambridge, MA, USA

<sup>6</sup>Center for Translational and Computational Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA

<sup>7</sup>Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA

<sup>8</sup>Division of Mental Health, Atlanta VA Medical Center, Decatur, GA, USA

<sup>9</sup>Department of Psychiatry, Emory University School of Medicine, Atlanta, GA, US

### Abstract

**Background:** Cognitive trajectory varies widely and can distinguish those that develop dementia from those that remain cognitively normal. Variation in cognitive trajectory is only partially explained by traditional neuropathologies. Here, we sought to identify novel genes associated with cognitive trajectory using DNA methylation profiles from human post-mortem brain.

**Methods:** We performed a brain epigenome-wide association study of cognitive trajectory in 636 participants from the Religious Order Study and the Rush Memory and Aging Project (ROS/MAP) using DNA methylation profiles of the dorsal lateral prefrontal cortex (dPFC). To maximize our

---

**Corresponding authors:** Michael P. Epstein, Ph.D., 305L Whitehead Building, 615 Michael Street NE, Atlanta, GA 30322-1047, mpepste@emory.edu, Phone: 404-712-8289; Thomas S. Wingo, M.D., 505K Whitehead Building, 615 Michael Street NE, Atlanta, GA 30322-1047, Thomas.wingo@emory.edu, Phone: 404-727-4905.  
#equal contribution

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of Interests

The authors report no biomedical financial interests or potential conflicts of interest.

power to detect epigenetic associations, we used the recently developed Gene Association with Multiple Traits (GAMuT) test to analyze the five measured cognitive domains simultaneously.

**Results:** We found an epigenome-wide association for differential methylation of sites in the Claudin-5 (*CLDN5*) locus and cognitive trajectory (p-value =  $9.96 \times 10^{-7}$ ), which was robust to adjustment for cell type proportions (p-value =  $8.52 \times 10^{-7}$ ). This association was primarily driven by association with declines in episodic (p-value =  $4.65 \times 10^{-6}$ ) and working memory (p-value =  $2.54 \times 10^{-7}$ ). This association between methylation in *CLDN5* and cognitive decline was even significant in participants with no or little signs of beta-amyloid and neurofibrillary tangle pathology.

**Conclusions:** Differential methylation of *CLDN5*, an important protein of the blood-brain barrier, is associated with cognitive trajectory beyond traditional Alzheimer Disease (AD) pathologies. The association between *CLDN5* and cognitive trajectory in people with low pathology suggests an early role for *CLDN5* and blood-brain barrier dysfunction in cognitive decline and AD.

### Keywords

Epigenetics; cognition; cognitive trajectory; dementia; gene-based analysis; neuropathology

## Introduction

Cognitive decline is a common concern among older adults; however, the trajectory of cognitive performance with age has a wide range from stable to rapid decline. Cognitive trajectory is an important predictor of health outcomes and mortality, independent of other commonly assessed risk factors (1). Dementia is a common consequence of a decline in cognition, and Alzheimer Disease (AD) is its leading cause (2). AD is characterized by the neuropathological accumulation of neuritic plaques and neurofibrillary tangles, which is accompanied by neuronal loss (3); however, most older individuals have several co-occurring neuropathologies. Collectively, neuropathologies explain about 40% of the variance in cognitive trajectory, leaving most unexplained (4,5). Thus, cognitive trajectory may be considered a summation of the different neuropathological and biological processes independent of pathologies at work in the aging human brain (5–7).

Despite the importance of understanding cognitive trajectory, existing epigenetic work on DNA methylation levels measured in brain tissue has primarily focused on AD-specific pathologies (8–11) and clinical diagnosis of AD (12–14). In contrast, epigenetic studies that focused on examining cognitive decline were limited due to measuring DNA methylation changes in blood (15), which showed only moderate correlations ( $\mu 0.4$ ) with brain methylation (15). Thus, there is need to understand the epigenetic changes that are associated with cognitive trajectory to identify potential mechanisms that may act through or independent of known neuropathologies.

In this study, we investigated associations between brain tissue-based DNA methylation and cognitive trajectory in 636 participants from the Religious Order Study and Rush Memory and Aging Project (ROS/MAP) cohorts. Cognitive trajectory was assessed in five

cognitive domains (episodic memory, perceptual speed, perceptual orientation, semantic memory, and working memory), which were analyzed simultaneously by using an innovative kernel procedure that can detect associations between multiple predictors (e.g. methylation sites in a gene) with multiple outcomes (e.g. multiple cognitive domains) (16,17), and identified genome-wide differential methylation of sites in the Claudin-5 (*CLDN5*) locus with cognitive trajectory (p-value =  $9.96 \times 10^{-7}$ ). Subsequent sensitivity analyses showed the differential methylation in the *CLDN5* locus was robust to cell type proportions, primarily driven by changes in episodic and working memory, and observed in individuals with little to no beta-amyloid and neurofibrillary tangle pathology.

## Materials and Methods

### Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP) Study Data

The Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP) are two large, prospectively followed cohorts (11,18). Both ROS and MAP recruit cognitively normal individuals and collect detailed annual cognitive and clinical evaluations. At death, participants donate their brain for detailed neuropathological assessments. Both studies were approved by an Institutional Review Board of Rush University Medical Center.

To be included in the present study, participants must have at least two follow-up evaluations and available brain methylation data derived from the dorsolateral prefrontal cortex (dPFC; Brodmann area 46). As in previous publications, the ROS and MAP data were analyzed jointly since much of the phenotypic data collected are identical at the item level in both studies and collected by the same investigative team (11,19). Cognitive trajectory was estimated using all available visits. Annual cognitive testing was converted to a z-score using the mean and standard deviation of the studies at baseline. Cognitive trajectory was modeled longitudinally using a mixed effects model, adjusting for age, sex, and education with a random slope and intercept. Neuropathologic measures of neuritic plaques and neurofibrillary tangles were made using Consortium to Establish a Registry for Alzheimer's Disease (CERAD) and Braak staging, respectively, using established research protocols (26).

### ROS/MAP Study Omics Data

The omics data generated on the ROS/MAP cohorts have been described previously. Each data source is presented here and complete details are presented in the supplementary materials.

DNA methylation was measured from the dPFC as previously described in 737 ROS/MAP participant samples (11), of which 665 had complete phenotype and covariate information. For this study, we re-processed the raw IDAT files, which were downloaded from Synapse (see Web Resources).

---

#### Web Resources

- Methylation data used in this study: <https://www.synapse.org/#!Synapse:syn3157275>
- Rush Alzheimer's Disease Center Research Resource Sharing Hub: <https://www.radc.rush.edu>
- Epstein Software: <https://github.com/epstein-software/GAMuT>.
- MRC London Brain Bank for Neurodegenerative Disease: <https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/mrc-london-neurodegenerative-diseases-brain-bank/>
- Ensembl gene predictions (ensGene, version of Apr-06-2014): <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/>.

Genotyping was generated using two microarrays, Affymetrix GeneChip 6.0 (Affymetrix, Inc, Santa Clara, CA, USA) and Illumina HumanOmniExpress (Illumina, Inc, San Diego, CA, USA), as described previously (20).

Gene expression was generated from the dPFC by Illumina HiSeq with 101-bp paired-end reads using the strand-specific dUTP method with poly-A selection with a coverage of 50 million reads. Single-cell RNA-sequencing profiled from the dPFC of 48 individuals from ROS/MAP cohort was performed as described (37).

### Statistical analysis

In our primary analysis, we estimated epigenetic associations across five neurocognitive domains in a gene-based analysis, in which each CpG site was assigned to the closest gene using the Bioconductor package hiAnnotator and the ensembl gene predictions (ensGene, version of Apr-06-2014). All CpG sites with a distance of no more than 20 KB to the closest gene, were included in the analyses. In addition, we conducted a sensitivity analysis in which we only included CpG sites with a distance of no more than 10 KB to the closest gene. In a traditional association study of cognitive trajectory, each cognitive test may be either tested individually (15) or used to estimate a composite measure aggregated across several cognitive tests (27). However, these approaches are underpowered in the presence of pleiotropy since they fail to exploit correlation among domains (17). Thus, analysis of a single composite measure can lose power if the causal CpG sites are only associated with a subset of the features that make the composite measure (17). Hence, it can be more powerful to directly account for the trait correlations using kernel methods (28). Kernel methods quantify the genetic similarity among pairs of subjects and test whether this genetic similarity is associated with trait similarity. Thus, they harness potential pleiotropy that exists between traits to improve power to detect associations. To analyze epigenetic associations across five neurocognitive domains simultaneously, we used the Gene Association with Multiple Traits (GAMuT) test (16). GAMuT is motivated by the idea that individuals with similar epigenetic patterns should also have similar cognitive traits across the different cognitive domains. Consequently, GAMuT constructs two different similarity matrices; one similarity matrix including cognitive decline in the five cognitive domains and the other similarity matrix for the epigenetic variation (beta values of CpG sites) assigned to a gene. Phenotypic and epigenetic similarity were modelled using linear kernels. P-values for GAMuT were derived using Davies' exact method, which is a computationally efficient method to provide accurate p-values in the extreme tails of tests that follow mixtures of chi-square variables (29). We applied a Bonferroni threshold to correct for multiple testing. The significance threshold was adjusted for the number of tested genes (threshold:  $0.05/26,558 = 1.88 \times 10^{-6}$ ).

In secondary analyses, we tested which cognitive domains and CpG sites were likely main drivers in our multivariate GAMuT analyses. The gene-based analyses for the single domains were performed with GAMuT and linear regression analyses were used in the CpG-based analyses.

All association models were adjusted for age at death, education, sex, ancestry, smoking status and post-mortem interval (PMI). Another important confounder in DNA methylation

analyses are differences in cell type proportions between study samples, which can be associated with diseases or phenotypes (30). In our main analyses, we used surrogate variables to adjust for differences in cell type proportions as recommended in reference (31). Principal components (PCs) based on CpG sites chosen for their potential to proxy nearby SNPs (within 10 BP) were used to correct simultaneously for cell type heterogeneity and population stratification (first three PCs, Figure S1) (32). Samples whose first PC (PC1) deviated more than 3 standard deviations from the mean PC1, were excluded from analyses, reducing the final sample size from 665 to 636. In a sensitivity analysis, associations were additionally adjusted for cell type proportions using a reference-based approach proposed in reference (33), which estimated the proportion of neurons from the measured DNA methylation beta values. To further validate the robustness of our main findings to the possibility of unmeasured confounding, we conducted a multivariate robust linear regression model with empirical Bayes from the R package limma (version 3.40.6) (34) controlling for unmeasured confounding using the eigenvalue difference method as implemented in the R package cate (35). All analyses were performed using R (version 3.4.3) using built-in functions unless otherwise specified.

To understand the results of the primary analysis in the ROS/MAP dataset, we performed the following additional analyses. First, we asked whether our primary finding is independent of traditional AD neuropathology, i.e., CERAD and Braak staging. This was achieved using interaction and mediation analyses. Mediation analysis was performed by using the R package “mediation”, an approach that relies on the quasi-Bayesian Monte Carlo method based on normal approximation (36), and enables us to determine the proportion of the association between methylation and cognitive decline that was mediated by neuropathology. Next, we assessed whether DNA methylation changes were associated with gene mRNA expression. Finally, we assessed whether the methylation signals we detected were the possible result of hidden genotype effects. All of these associations were tested using GAMuT and the genotype associations were followed by a linear regression analysis on the single SNP level.

To test for generalizability of our findings from the ROS/MAP dataset, we examined whether the DNA methylation signals we found associated with cognitive decline and Braak staging were associated with Braak staging in MRC London Brain Bank for Neurodegenerative Disease dataset (GSE59685, see supplementary methods for details; Braak stage was the only outcome available in this dataset).

## Results

### Description of study participants

Our study sample consisted of 636 individuals from the ROS/MAP cohorts with an average age at death of 86 years and with 63% being female (Table 1). Most of the participants were white (98%), had a high level of education and 70% had never smoked. On average, cognitive performance declined with age for every single domain (Table 1) and correlations of cognitive decline between different domains were moderate, ranging from 0.54 to 0.78 (Table S1). Based on their neuritic plaques (CERAD score), 35% had no or possible AD at time of death and 65% of the participants had definite or probable AD. Neurofibrillary

tangle (NFT) pathology was evaluated using the Braak stage. Most participants had a Braak stage of III (30%) or higher (IV: 28%, V: 22%, VI: 1%), which indicates involvement of limbic regions (stages III and IV) or moderate to severe neocortical involvement (V and VI). Cognitive decline was associated with more signs of neuropathology (CERAD and Braak stage, Table S2).

### Methylation patterns of *CLDN5* associated with cognitive decline

We found that methylated CpG sites in the Claudin-5 (*CLDN5*) locus were associated with cognitive trajectory (p-value =  $9.96 \times 10^{-7}$ ; Figure 1, Table 2, Table S3, and Figure S1). This association was robust to adjustment for cell type proportions (p-value =  $8.52 \times 10^{-7}$ , Table S4, Figures S2 and S3), to the selection of a smaller window of CpG sites around each gene (p-value =  $9.96 \times 10^{-7}$ , within 10kb, Table S5), and to the restriction to participants with European ancestry (621/636 participants, p-value =  $9.73 \times 10^{-7}$ , Table S6). The trajectories of episodic and working memory were the main drivers for the observed association with both being associated with methylation in the *CLDN5* locus in the analyses of the single domains (Table 2, Figure S4 and Tables S7–S11). Genes showing suggestive association with cognitive trajectory (p-values  $< 5 \times 10^{-5}$ ) included *AC084018.1*, *CTB-186G2.1*, *ATG16L2*, *KCNN4*, *RP11-779O18.1*, *TTC22*, *DCUN1D2-AS*, *PNMA1* and *RP11-101C11.1*. The strongest associations with these genes were found with episodic memory, followed by working memory. Interestingly, most of these methylation signals (*CLDN5* and 7/9 suggestive genes) were also at least nominally associated with CERAD (Figures S5–S6, Tables S12 and S13) and Braak stage (Figure S7, Tables S12 and S14).

### Higher levels of methylation in *CLDN5* locus associated with cognitive decline

To identify which CpG sites are the main drivers of the observed associations and to understand the direction of association, we conducted a linear regression analysis for the cognitive trajectory of each cognitive domain. Interestingly, except for *PNMA1*, higher levels of methylation within our top genes were associated with an increased cognitive decline in every single cognitive domain (Table S15). Within the CpG sites assigned to *CLDN5*, cg16773741 and cg05460329 were the main drivers of the association with cg16773741 being associated with episodic memory (p-value =  $1.48 \times 10^{-8}$ ), semantic memory ( $8.81 \times 10^{-8}$ ) and working memory ( $8.66 \times 10^{-9}$ ) (Table 3A). CpG sites located in close proximity to these two CpG sites showed weaker associations with cognitive trajectory, but were still at least nominally significant (Table 3, Figure S9). This indicates that methylation in the whole epigenomic region seems to be related to cognitive trajectory. These associations were robust to additional adjustment for unmeasured confounding (Table 3B).

### Association with *CLDN5* even present without signs of neuropathology

To investigate if associations between methylation in *CLDN5* and cognitive decline are also present in participants without clear signs of neuropathology, we conducted an analysis of the interaction between the most significantly associated CpG site (cg16773741) and CERAD or Braak stage for cognitive decline. The association between methylation in cg16773741 did not significantly differ between participants with no to little signs of neuropathology versus participants with moderate to severe signs of neuropathology

(measured by CERAD and Braak stage; Figure 2). Consequently, the association between methylation in cg16773741 and decline in episodic, semantic, and working memory was even significant in participants with no or little signs of neuropathology.

### Partial mediation through neuropathology

The association between DNA methylation in the *CLDN5* locus (cg16773741) and cognitive trajectory was only partially mediated through an increased neuropathology. It ranged between 17% (95% confidence interval (CI): 18–40%) to 31% (95% CI: 18–61%) for CERAD and 13% (95% CI: 7–21%) to 27% (95% CI: 15–41%) for Braak stage depending on the cognitive domain (Figure 3). Therefore, the major part of the association with *CLDN5* (total effect, Figure 3) was a direct association between methylation and cognitive decline, which was independent of beta-amyloid and neurofibrillary neuropathology (direct effect, Figure 3).

### No clear association with gene expression levels

In our sample, we found no association between DNA methylation in the *CLDN5* window and *CLDN5* expression (p-value = 0.1978; Table S16). *KCNN4* was the only top gene (Table 2) for which methylation levels were associated with expression levels (p-value = 0.0004; Table S16). Furthermore, cognitive trajectory was not associated with expression of any gene in Table 2 (Table S16).

### Methylation signals were independent of genotypes

Genotypes located within the *CLDN5* locus were not associated with cognitive trajectory (p-value = 0.4415, Table S17 A; Table S18) and associations between DNA methylation in the *CLDN5* locus and cognitive trajectory were robust to adjustment for genotypes in the same locus (Table S17 B).

### Generalizability of our findings from the ROS/MAP dataset

To test for generalizability of our findings from the ROS/MAP dataset, we examined whether the DNA methylation signals we found associated with cognitive decline and Braak staging in ROS/MAP were associated with Braak staging in participants of the MRC London Brain Bank for Neurodegenerative Disease. Compared to ROS/MAP, individuals from the brain bank tended to have a higher mean Braak staging (4.7, SD=1.6), indicating a higher degree of neurofibrillary tangle burden but have a similar mean age at death (87 years, SD = 7 year) and proportion of women (67%).

For the 8 methylation signals, which showed suggestive association with cognitive trajectory (p-values  $< 5 \times 10^{-5}$ ) and were at least nominally associated with Braak stage in ROS/MAP (p-values  $< 0.05$ ), we observed nominal associations of Braak stage with *CLDN5*, *CTB-186G2.1* and *KCNN4* loci (Table S12, Figure S8) in this additional dataset and the direction of association with the most significant CpG sites assigned to these genes was consistent with the association observed in ROS/MAP (Table S15, N=66).

### Single-cell RNA-sequencing shows enrichment of *CLDN5* in brain endothelial cells

To understand the specificity of our findings, we tested whether the top 10 signals were enriched in a particular brain cell-type. We obtained single-cells RNA-sequencing that was generated from dPFC from participants in ROS/MAP (37) and found that 3 of the 10, including the only genome-wide significantly associated region, *CLDN5*, showed evidence of significant association after FDR adjustment for all genes within the experiment (~17,000) (Table S19, Figure S10). Notably, *CLDN5*, was estimated to be 3.4-fold enriched in endothelial cells (FDR p-value  $2.8 \times 10^{-235}$ ), and *PNMA1* and *ATG16L2* were enriched 0.32 (FDR p-value  $1.3 \times 10^{-15}$ ) and 0.73 (FDR p-value  $6.7 \times 10^{-10}$ ) in excitatory neurons and microglia, respectively.

### Discussion

In this study, we found an epigenome-wide association between brain-tissue-based DNA methylation in the *CLDN5* locus and cognitive trajectory in more than 600 participants from the ROS/MAP cohort. This association was significant across different domains and particularly associated with trajectories in episodic and working memory. We also found that higher levels of methylation in *CLDN5* were associated with neuropathology in ROS/MAP and in an independent dataset consistent with the direction of association found with cognitive decline. Most interestingly, the association between methylation in *CLDN5* and cognitive decline was even present in participants with low levels of beta-amyloid and neurofibrillary neuropathology. In addition, only 13–31% of the association between methylation and cognitive decline was mediated through levels of neuropathology, whereas the major part of the association was independent of it. Finally, we found no evidence that hidden effects of genotypes in the *CLDN5* locus confounded our methylation results.

*CLDN5* is an integral membrane protein and an important component of tight junction protein complexes that comprise the blood-brain barrier. The blood-brain barrier is located at endothelial cells lining the brain microvasculature and is maintained by the neurovascular unit, a functional relationship between astrocytes, neurons, and endothelial cells (39). Dysfunction of the blood-brain barrier has been implicated in neurodegenerative disorders, such as AD (39–44), and depression (45). Prior studies examining *CLDN5* in neurodegeneration have primarily focused on its role in animal and *in vitro* models of AD (46,47), which are based on beta-amyloid pathology, or neuropathology in less than 50 human brain samples (48). Here, for the first time, we provide evidence from a large population-based cohort study of human brain implicating *CLDN5* with higher burden of neuropathology, consistent with prior studies (46). Remarkably, associations with *CLDN5* are present in individuals with low levels of neuropathology as well. Taken together with results from the previous studies, our results suggest that *CLDN5* is associated with cognitive trajectory beyond its effect on beta-amyloid. In total, the genetic changes we observe and the effect of *CLDN5* on cognitive trajectory in people with low pathology suggest an early role for *CLDN5* and blood-brain barrier dysfunction in cognitive decline and AD. We note that an estimated two-thirds of AD dementia (clinically defined) and that an estimated 40% of cognitive decline are attributable to known age-related



neuropathologies (4,49). Thus, our findings may account for some of the unaccounted-for variation in cognitive decline and AD dementia.

To the best of our knowledge, this is the first epigenome-wide study using cognitive trajectory in older individuals. A previous study on the same cohort showed an association of 71 CpG sites with neuritic plaque burden, of which 11 were validated in an independent cohort (11). Here, we showed that all of these 11 signals were at least nominally associated with cognitive trajectory (p-values < 0.05) and the strongest associations were again found for cognitive trajectory of episodic and working memory (Table S20). In addition, we identified methylation in *CLDN5* as new epigenetic factor associated with cognitive trajectory, a gene that has not been linked to AD in a population-based cohort.

By contrast, associations of blood-based methylation levels with global cognitive function (cg21450381) and phonemic verbal fluency (cg12507869) (15) could not be validated in our study (Table S21). The likely reasons are the different source of methylation data, and differences in the phenotype (i.e., cognitive trajectory over time in our study versus cognitive testing at a single time point (15)).

Strengths of this study include the use of the ROS/MAP cohort in our main analyses, which is unique in terms of its longitudinal nature with very high follow-up rates, prospective collection of data, a community-based cohort design, and detailed neuropathological examination following high autopsy rates. This study is also strengthened by the GAMuT (16) statistical method that harnesses correlations among cognitive domains and among CpG sites. Using GAMuT in contrast to traditional univariate single CpG analyses has the following advantages: 1) GAMuT leverages information from multiple correlated phenotypes, in this case cognitive trajectory across different domains, to improve power of discovery relative to univariate analysis of each separate domain; 2) GAMuT clusters DNA methylation array data by exploiting the biological nature of DNA methylation where certain regions of proximally located DNAm sites exhibit correlated methylation state (50). CpGs in close proximity to each other have been shown to typically function together as a unit (51) and have been investigated in terms of their correlated methylation status and whether they are joined together by various biological mechanisms. While <10% of the ~28 million CpGs in the human genome are in CpG dense regions, they are enriched in the promoters and transcription start sites of developmental and housekeeping genes, implying potential biological relevance (51). While cg16773741 was the lead signal for most associations, CpG sites located in close proximity to cg16773741 showed weaker associations with cognitive trajectory, but were still at least nominally significant. This indicates that methylation in the region is associated with cognitive trajectory, which increases the likelihood of a potential biological relevance of this finding.

Due to the unique study design of the ROS/MAP cohorts with longitudinal data on cognitive trajectory and neuropathologies and DNA methylation measured in brain tissue after death, we were not able to validate our findings in an independent dataset. The only available dataset was from a brain bank and had only data on the Braak stage. In addition, the study is potentially limited by the cross-sectional measurement of DNA methylation after death. Although brain tissue is the ideal target tissue to measure DNA methylation related

to cognitive trajectory, it inhibited a simultaneous (or even later) assessment of cognitive function. Consequently, we cannot exclude the potential risk of reverse causality in our associations. Another potential limitation is the use of bulk tissue analysis which does not allow fine-grain resolution of the influences of different cell populations, especially rare cell types like endothelial cells. This problem is partially mitigated by adjusting for cell-type composition; however, the mitigation approach works best for effects from common brain cell types (e.g., glia and neurons). Here, we used brain single-cell data to show *CLDN5* is exclusively enriched in brain endothelial cells, which agrees with prior human and mouse data showing that show *CLDN5* is a key protein for maintaining the integrity of the blood brain barrier (52). Thus, future studies should investigate the role of epigenetic changes we observed in the *CLDN5* locus in brain endothelial cells to elucidate whether there is a causal relationship between *CLDN5* gene regulation and cognitive decline. Finally, the Illumina 450k array only covers 1.7% of CpG sites on the human genome, and most of these are located in promoter regions. We therefore encourage more studies on this topic using more advanced EWAS platforms (such as the Illumina EPIC array) or sequencing approaches to further disentangle the epigenetic patterns of cognitive trajectory and neuropathology.

In conclusion, we have presented evidence for brain-based DNA methylation in association with cognitive trajectory. We identified methylation in *CLDN5* as a new epigenetic factor associated with cognitive trajectory, which was even significant in participants with no or little signs of beta-amyloid and neurofibrillary neuropathology. Higher levels of methylation in *CLDN5* were associated with cognitive decline implicating the blood brain barrier in maintenance of cognitive trajectory with aging.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors are grateful to the participants of the Rush Memory and Aging Project and Religious Orders Study and the Medical Research Counsel Brain Bank.

AH was supported by a research fellowship from the Deutsche Forschungsgemeinschaft (DFG; HU 2731/1–1) and by the HERCULES Center (NIEHS P30ES019776). MPE was supported by NIH grants R01 GM117946 and R01 AG071170. APW is supported by NIH grants R01 AG056533, VA I01 BX003853, and NIH U01 MH115484. TSW was supported by NIH grants P50 AG025688, R56 AG062256, R56 AG060757, and R01 AG056533. CR was supported by NIH grant T32 NS007480. DAB was supported by P30AG10161, R01AG15819, R01AG17917, R01AG36042, U01AG61356. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of manuscript. This article was published as a preprint on bioRxiv: doi: <https://doi.org/10.1101/857953>.

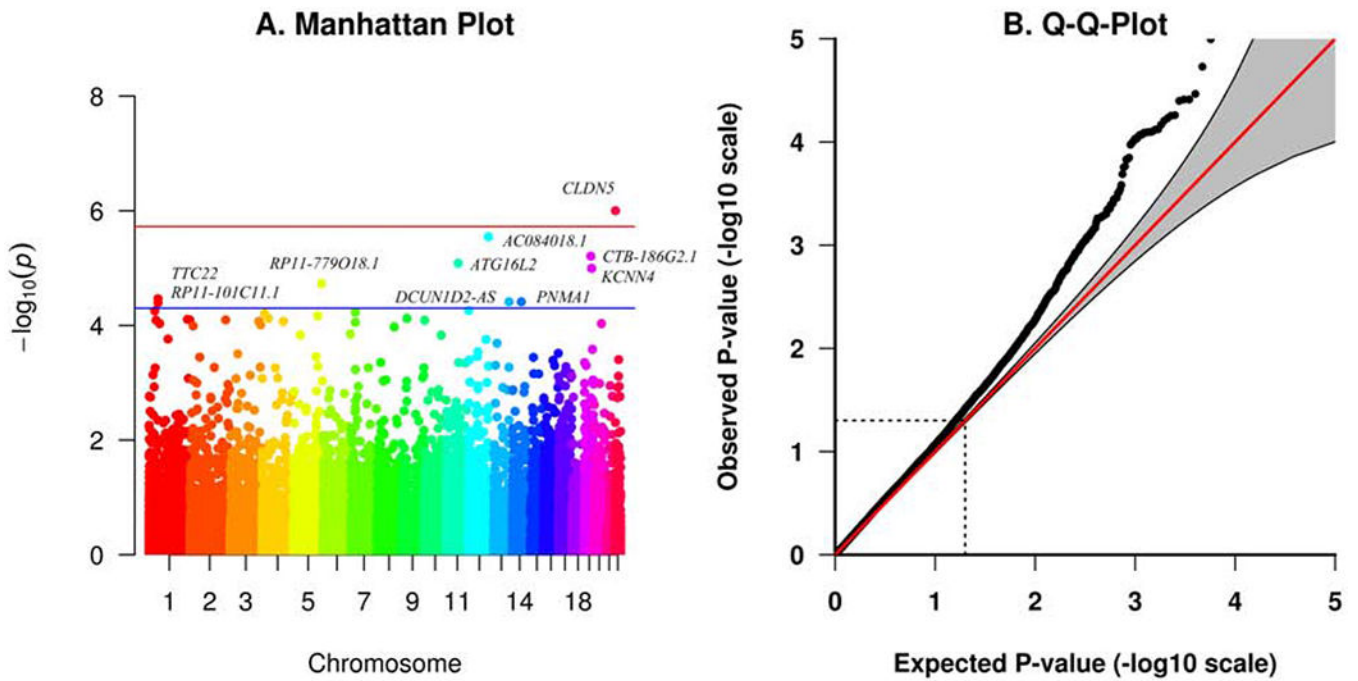
## References

1. Batty GD, Deary IJ, Zaninotto P (2016): Association of cognitive function with cause-specific mortality in middle and older age: Follow-up of participants in the english longitudinal study of ageing. *Am J Epidemiol* 183: 183–190. [PubMed: 26803665]
2. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K (2016): The genetic landscape of Alzheimer disease: Clinical implications and perspectives. *Genet Med* 18: 421–430. [PubMed: 26312828]

3. Möller HJ, Graeber MB (1998): The case described by Alois Alzheimer in 1911. Historical and conceptual perspectives based on the clinical record and neurohistological sections. *Eur Arch Psychiatry Clin Neurosci* 248: 111–122. [PubMed: 9728729]
4. Boyle PA, Wilson RS, Yu L, Barr AM, Honer WG, Schneider JA, Bennett DA (2013): Much of late life cognitive decline is not due to common neurodegenerative pathologies. *Ann Neurol* 74: 478–489. [PubMed: 23798485]
5. Boyle PA, Yu L, Wilson RS, Leurgans SE, Schneider JA, Bennett DA (2018): Person-specific contribution of neuropathologies to cognitive loss in old age. *Ann Neurol* 83: 74–83. [PubMed: 29244218]
6. Schneider JA, Aggarwal NT, Barnes L, Boyle P, Bennett DA (2009): The neuropathology of older persons with and without dementia from community versus clinic cohorts. *J Alzheimer's Dis* 18: 691–701. [PubMed: 19749406]
7. Terry RD, Masliah E, Salmon DP, Butters N, DeTeresa R, Hill R, et al. (1991): Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann Neurol* 30: 572–580. [PubMed: 1789684]
8. Li P, Marshall L, Oh G, Jakubowski JL, Groot D, He Y, et al. (2019): Epigenetic dysregulation of enhancers in neurons is associated with Alzheimer's disease pathology and cognitive symptoms. *Nat Commun* 10: 1–14. [PubMed: 30602773]
9. Altuna M, Urdániz-Casado A, Sánchez-Ruiz De Gordo J, Zelaya M V, Labarga A, Lepsant MJJ, et al. (2019): DNA methylation signature of human hippocampus in Alzheimer's disease is linked to neurogenesis. *Clin Epigenetics* 11: 1–16. [PubMed: 30611298]
10. Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. (2014): Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci* 17: 1164–1170. [PubMed: 25129077]
11. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. (2014): Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci* 17: 1156–1163. [PubMed: 25129075]
12. Humphries CE, Kohli MA, Nathanson L, Whitehead P, Beecham G, Martin E, et al. (2015): Integrated whole transcriptome and DNA methylation analysis identifies gene networks specific to late-onset Alzheimer's disease. *J Alzheimer's Dis* 44: 977–987. [PubMed: 25380588]
13. Watson CT, Roussos P, Garg P, Ho DJ, Azam N, Katsel PL, et al. (2016): Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Med* 8: 1–14. [PubMed: 26750923]
14. Bakulski KM, Dolinoy DC, Sartor MA, Paulson HL, Konen JR, Lieberman AP, et al. (2012): Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. *J Alzheimer's Dis* 29: 571–588. [PubMed: 22451312]
15. Marioni RE, McRae AF, Bressler J, Colicino E, Hannon E, Li S, et al. (2018): Meta-analysis of epigenome-wide association studies of cognitive abilities. *Mol Psychiatry* 23: 2133–2144. [PubMed: 29311653]
16. Broadaway KA, Cutler DJ, Duncan R, Moore JL, Ware EB, Jhun MA, et al. (2016): A statistical approach for testing cross-phenotype effects of rare variants. [PMC4800053]. *Am J Hum Genet* 98: 525–540. [PubMed: 26942286]
17. Holleman AM, Broadaway KA, Duncan R, Todor A, Almlí LM, Bradley B, et al. (2019): Powerful and Efficient Strategies for Genetic Association Testing of Symptom and Questionnaire Data in Psychiatric Genetic Studies. *Sci Rep* 9: 1–11. [PubMed: 30626917]
18. Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA (2018): Religious Orders Study and Rush Memory and Aging Project. *J Alzheimer's Dis* 64: S161–S189. [PubMed: 29865057]
19. Bennett DA, Wilson RS, Boyle PA, Buchman AS, Schneider JA (2012): Relation of neuropathology to cognition in persons without cognitive impairment. *Ann Neurol* 72: 599–609. [PubMed: 23109154]

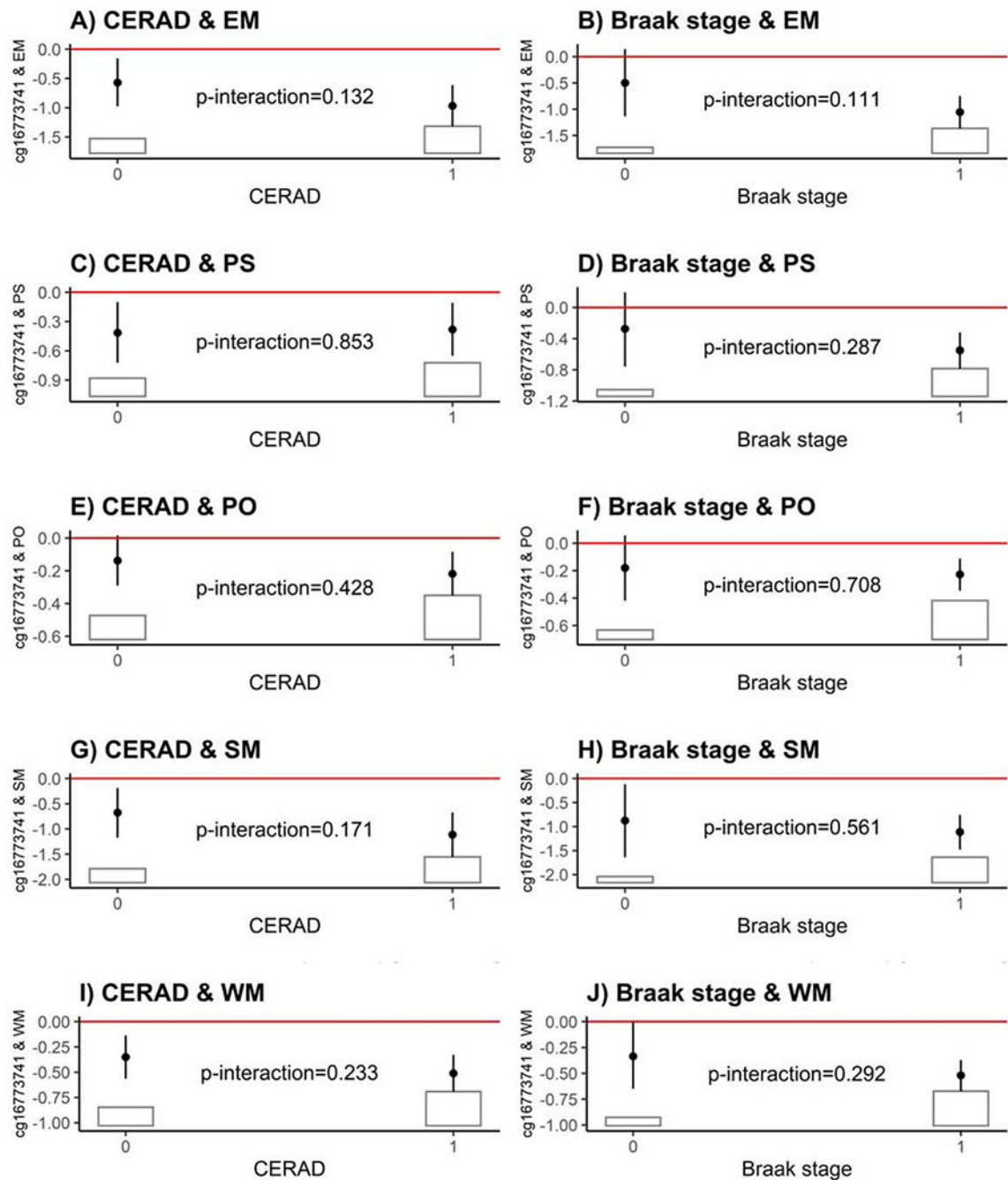
20. De Jager PL, Shulman JM, Chibnik LB, Keenan BT, Raj T, Wilson RS, et al. (2012): A genome-wide scan for common variants affecting the rate of age-related cognitive decline. *Neurobiol Aging* 33: 1017.e1–15.
21. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. (2016): Next-generation genotype imputation service and methods. *Nat Genet* 48: 1284–1287. [PubMed: 27571263]
22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2013): STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21. [PubMed: 23104886]
23. Love MI, Huber W, Anders S (2014): Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550. [PubMed: 25516281]
24. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. (2015): A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A* 112: 7285–7290. [PubMed: 26060301]
25. Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS (2012): Overview and findings from the religious orders study. *Curr Alzheimer Res* 9: 628–645. [PubMed: 22471860]
26. Braak H, Braak E (1991): Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 82: 239–259. [PubMed: 1759558]
27. Paajanen T, Hänninen T, Tunnard C, Hallikainen M, Mecocci P, Sobow T, et al. (2014): CERAD neuropsychological compound scores are accurate in detecting prodromal alzheimer's disease: A prospective AddNeuroMed study. *J Alzheimer's Dis* 39: 679–690. [PubMed: 24246420]
28. Larson NB, Chen J, Schaid DJ (2019): A review of kernel methods for genetic association studies. *Genet Epidemiol* 43: 122–136. [PubMed: 30604442]
29. Davies RB (1980): Algorithm AS 155: The Distribution of a Linear Combination of  $\chi^2$  Random Variables. *J R Stat Soc Ser C Appl Stat* 29: 323–333.
30. Laird PW (2010): Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet* 11. 10.1038/nrg2732
31. Mcgregor K, Bernatsky S, Colmegna I, Hudson M, Pastinen T, Labbe A, Greenwood CMT (2016): An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. *Genome Biol* 1–17. [PubMed: 26753840]
32. Barfield RT, Almi LM, Kilaru V, Smith AK, Mercer KB, Duncan R, et al. (2014): Accounting for population stratification in DNA methylation studies. *Genet Epidemiol* 38: 231–241. [PubMed: 24478250]
33. Guintivano J, Aryee MJ, Kaminsky ZA (2013): A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics* 8: 290–302. [PubMed: 23426267]
34. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47. [PubMed: 25605792]
35. Wang J, Zhao Q (2019): cate: High Dimensional Factor Analysis and Confounder Adjusted Testing [no. R package version 1.1]. Retrieved from <https://cran.r-project.org/package=cate>
36. Imai K, Keele L, Tingley D (2010): A general approach to causal mediation analysis. *Psychol Methods* 15: 309–334. [PubMed: 20954780]
37. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. (2019): Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 570: 332–337. [PubMed: 31042697]
38. Wingo A, Liu Y, Gerasimov E, Gockley J, Logsdon B, Duong D, et al. (2021): Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet* in press.
39. Marques F, Sousa JC, Sousa N, Palha JA (2013): Blood-brain-barriers in aging and in Alzheimer's disease. *Mol Neurodegener* 8: 38. [PubMed: 24148264]
40. Farrall AJ, Wardlaw JM (2009): Blood-brain barrier: ageing and microvascular disease--systematic review and meta-analysis. *Neurobiol Aging* 30: 337–352. [PubMed: 17869382]
41. Viggars AP, Wharton SB, Simpson JE, Matthews FE, Brayne C, Savva GM, et al. (2011): Alterations in the blood brain barrier in ageing cerebral cortex in relationship to Alzheimer-type pathology: a study in the MRC-CFAS population neuropathology cohort. *Neurosci Lett* 505: 25–30. [PubMed: 21970975]

42. Weiss N, Miller F, Cazaubon S, Couraud P-O (2009): The blood-brain barrier in brain homeostasis and neurological diseases. *Biochim Biophys Acta* 1788: 842–857. [PubMed: 19061857]
43. Montagne A, Barnes SR, Sweeney MD, Halliday MR, Sagare AP, Zhao Z, et al. (2015): Blood-brain barrier breakdown in the aging human hippocampus. *Neuron* 85: 296–302. [PubMed: 25611508]
44. Nation DA, Sweeney MD, Montagne A, Sagare AP, D’Orazio LM, Pachicano M, et al. (2019): Blood-brain barrier breakdown is an early biomarker of human cognitive dysfunction. *Nat Med* 25: 270–276. [PubMed: 30643288]
45. Dudek KA, Dion-Albert L, Lebel M, LeClair K, Labrecque S, Tuck E, et al. (2020): Molecular adaptations of the blood-brain barrier promote stress resilience vs. depression. *Proc Natl Acad Sci U S A* 117: 3326–3336. [PubMed: 31974313]
46. Keaney J, Walsh DM, O’Malley T, Hudson N, Crosbie DE, Loftus T, et al. (2015): Autoregulated paracellular clearance of amyloid- $\beta$  across the blood-brain barrier. *Sci Adv* 1: e1500472. [PubMed: 26491725]
47. Shimizu F, Sano Y, Saito K, Abe M, Maeda T, Haruki H, Kanda T (2012): Pericyte-derived glial cell line-derived neurotrophic factor increase the expression of claudin-5 in the blood-brain barrier and the blood-nerve barrier. *Neurochem Res* 37: 401–409. [PubMed: 22002662]
48. Yamazaki Y, Shinohara M, Shinohara M, Yamazaki A, Murray ME, Liesinger AM, et al. (2019): Selective loss of cortical endothelial tight junction proteins during Alzheimer’s disease progression. *Brain* 142: 1077–1092. [PubMed: 30770921]
49. Boyle PA, Yu L, Leurgans SE, Wilson RS, Brookmeyer R, Schneider JA, Bennett DA (2019): Attributable risk of Alzheimer’s dementia attributed to age-related neuropathologies. *Ann Neurol* 85: 114–124. [PubMed: 30421454]
50. Gatev E, Gladish N, Mostafavi S, Kobor MS (2020): CoMeBack: DNA methylation array data analysis for co-methylated regions. *Bioinformatics* 36: 2675–2683. [PubMed: 31985744]
51. Deaton AM, Bird A (2011): CpG islands and the regulation of transcription. *Genes Dev* 25:1010–1022. [PubMed: 21576262]
52. Daneman R, Prat A (2015): The Blood-Brain Barrier. *Cold Spring Harb Perspect Biol* 7: a020412. [PubMed: 25561720]



**Figure 1. DNA methylation and cognitive decline.**

Association between DNA methylation and cognitive decline in ROS/MAP tested with GAMuT. Adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first three principal components.

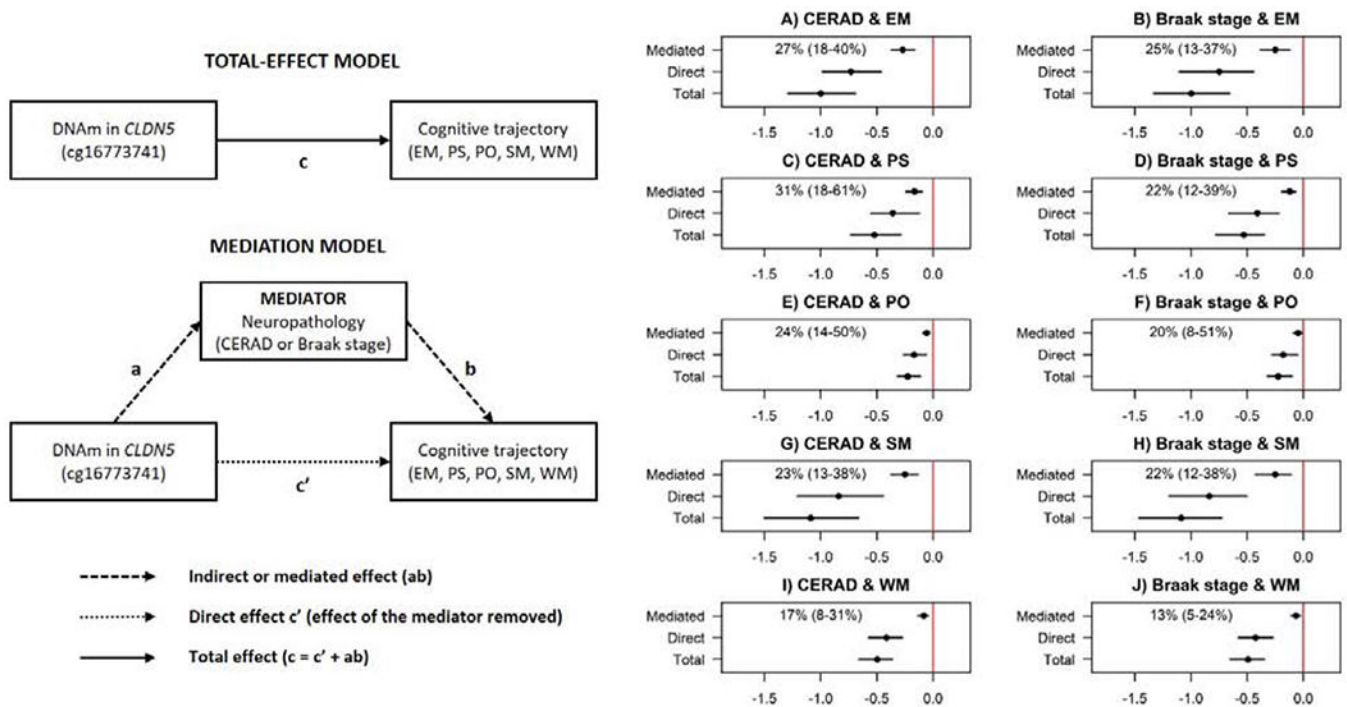


**Figure 2. Interaction analysis.**

Associations between DNA methylation of the top CpG site of *CLDN5* (cg16773741) and cognitive trajectory are shown in participants with no to little (category 0) vs. moderate to severe (category 1) signs of neuropathology (beta estimates and 95%-confidence intervals). No to little signs of neuropathology are defined as a CERAD measure of 3 (possible) or 4 (no AD) or a Braak stage of 0 to II. Moderate to severe signs of neuropathology are defined as a CERAD measure of 1 (definite) or 2 (probable) or a Braak stage of III to VI. P-values are given for the test of deviations of the association between methylation

in cognitive trajectory between the two strata. The boxes present the distribution of the neuropathological variables (proportion of samples in category 0 and 1). EM: decline in episodic memory; PS: decline in perceptual speed; PO: decline in perceptual orientation; SM: decline in semantic memory; WM: decline in working memory. Adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first three principal components.





**Figure 3. Mediation analysis.**

Beta-estimates and 95%-confidence intervals of the estimated average causal mediation effects, the average direct effects as well as the total effects. Proportion (with 95%-confidence interval) of the association between DNA methylation (DNAm) of the top CpG site of *CLDN5* (cg16773741) and cognitive trajectory, which is mediated through neuropathology (CERAD & Braak stage) is given in percent. EM: decline in episodic memory; PS: decline in perceptual speed; PO: decline in perceptual orientation; SM: decline in semantic memory; WM: decline in working memory. Adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first three principal components (PCs).

**Table 1.**

## Study characteristics.

	ROS/MAP
N	636
Age at death, mean $\pm$ sd	86.22 $\pm$ 4.72
Female, n (%)	401 (63.1)
Ancestry	
European, n (%)	621 (97.6)
African American, n (%)	11 (1.7)
Native American, n (%)	1 (0.2)
Asian, n (%)	3 (0.5)
Years of education, mean $\pm$ sd	16.63 $\pm$ 3.54
Never smoker, n (%)	444 (69.8)
Ex-smoker, n (%)	176 (27.7)
Smoker, n (%)	16 (2.5)
Post mortem interval (PMI), mean $\pm$ sd	7.43 $\pm$ 5.79
Estimated proportion of neurons <sup>1</sup> , mean $\pm$ sd	0.45 $\pm$ 0.05
Decline in episodic memory, mean $\pm$ sd	-0.03 $\pm$ 0.11
Decline in perceptual speed, mean $\pm$ sd	-0.02 $\pm$ 0.08
Decline in perceptual orientation, mean $\pm$ sd	-0.01 $\pm$ 0.04
Decline in semantic memory, mean $\pm$ sd	-0.03 $\pm$ 0.13
Decline in working memory, mean $\pm$ sd	-0.01 $\pm$ 0.05
CERAD score <sup>2</sup> , n (%)	
Definite	204 (32.1)
Probable	208 (32.7)
Possible	70 (11.0)
No Alzheimer's disease	154 (24.2)
Braak stage <sup>3</sup> , n (%)	
0	7 (1.1)
I	56 (8.8)
II	61 (9.6)
III	188 (29.6)
IV	180 (28.3)
V	138 (21.7)
VI	6 (0.9)

<sup>1</sup>in comparison to glia cells. Estimated proportion of neurons is highly correlated with PC2 (Pearson correlation  $r=-0.5$ ) and PC3 ( $r=0.8$ ). Correlations with PC1 and PC4 are  $r=-0.1$ .

<sup>2</sup>Semiquantitative measure of neuritic plaques;

<sup>3</sup>Semiquantitative measure of neurofibrillary tangle (NFT) pathology. Braak stages I and II indicate NFTs confined mainly to the entorhinal region of the brain, Braak stages III and IV indicate involvement of limbic regions such as the hippocampus, Braak stages V and VI indicate moderate to severe neocortical involvement.

**Table 2.**

Top signals (p-values < 5 x 10<sup>-5</sup>) for the association between DNA methylation and cognitive decline across all domains as well as in the single domains in ROS/MAP.

gene	chr	pos (lower)	pos (upper)	mean dist gene	max dist gene	#CpG sites	p-value (across all domains)	p-value (EM)	p-value (PS)	p-value (PO)	p-value (SM)	p-value (WM)
CLDN5	chr22	19510877	19515608	1030	2062	18	<b>9.96E-07</b>	4.65E-06	7.16E-05	0.0021	<u>2.35E-05</u>	<b>2.54E-07</b>
AC084018.1	chr12	122235169	122241475	501	1414	15	<u>2.84E-06</u>	5.96E-05	0.0001	0.0060	<u>4.13E-06</u>	0.0002
CTB-186G2.1	chr19	39087135	39090701	692	1633	4	<u>6.19E-06</u>	<u>7.33E-06</u>	<u>3.50E-05</u>	0.0015	7.45E-05	7.69E-05
ATG16L2	chr11	72521478	72546168	1326	12094	25	<u>8.22E-06</u>	<u>1.32E-05</u>	0.0002	0.0086	8.31E-05	0.0004
KCNN4	chr19	44270892	44286076	1682	6854	13	<u>1.01E-05</u>	<u>2.01E-05</u>	0.0002	0.0127	0.0001	<u>4.12E-05</u>
RP11-779O18.1	chr5	172168177	172189374	7685	17051	12	<u>1.86E-05</u>	<u>1.26E-05</u>	0.0002	0.0262	0.0009	<u>3.18E-06</u>
TTC22	chr1	55240609	55268659	1189	4775	23	<u>3.42E-05</u>	0.000192	0.0002	0.0182	0.0002	6.16E-05
DCUNID2-AS	chr13	114123258	114129580	2530	3583	10	<u>3.87E-05</u>	<u>5.91E-06</u>	0.0004	0.0176	0.0014	<u>4.48E-05</u>
PNMA1	chr14	74177136	74181427	777	1357	9	<u>3.88E-05</u>	<u>3.14E-05</u>	0.0079	0.0030	0.0002	0.0006
RP11-101C11.1	chr1	55682652	55709508	5270	9658	2	<u>3.99E-05</u>	<u>2.36E-05</u>	0.0004	0.0565	0.0006	0.0003

EM: decline in episodic memory; PS: decline in perceptual speed; PO: decline in perceptual orientation; SM: decline in semantic memory; WM: decline in working memory  
 Adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first three principal components (PCs).

Bonferroni threshold: 0.05/26,558=1.88x10<sup>-6</sup> (p-values below Bonferroni threshold in **bold**)

Suggestive: p-values < 5x10<sup>-5</sup> (underlined)

**Table 3.**

Identification of lead signals within *CLDN5* and direction of associations.

A. Linear regression adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first four PCs													
gene	chr	pos	CpG site	beta (EM)	p-value (EM)	beta (PS)	p-value (PS)	beta (PO)	p-value (PO)	beta (SM)	p-value (SM)	beta (WM)	p-value (WM)
CLDN5	chr22	19512903	cg05460329	-0.93	1.43E-06	-0.62	1.60E-05	-0.25	0.0003	-1.10	1.89E-06	-0.45	2.41E-06
CLDN5	chr22	19512942	cg05498726	-0.55	0.0001	-0.36	0.0008	-0.13	0.0162	-0.61	0.0005	-0.26	0.0002
CLDN5	chr22	19513006	cg11450827	-0.60	0.0007	-0.35	0.0064	-0.16	0.0111	-0.65	0.0019	-0.30	0.0005
CLDN5	chr22	19513008	cg17583256	-0.57	0.0005	-0.30	0.0129	-0.14	0.0180	-0.51	0.0086	-0.27	0.0008
CLDN5	chr22	19513017	cg16773741	-0.76	<b>1.48E-08</b>	-0.39	8.38E-05	-0.18	0.0003	-0.86	<b>8.81E-08</b>	-0.38	<b>8.66E-09</b>
CLDN5	chr22	19513078	cg14553765	-0.44	0.0014	-0.32	0.0017	-0.13	0.0087	-0.60	0.0003	-0.29	1.92E-05
CLDN5	chr22	19513176	cg00189989	-0.61	0.0004	-0.42	0.0009	-0.14	0.0224	-0.78	0.0001	-0.36	1.86E-05

B. Multivariate robust linear regression additionally adjusted for unmeasured confounding using the eigenvalue difference method <sup>1</sup>													
gene	chr	pos	CpG site	beta (EM)	p-value (EM)	beta (PS)	p-value (PS)	beta (PO)	p-value (PO)	beta (SM)	p-value (SM)	beta (WM)	p-value (WM)
CLDN5	chr22	19512903	cg05460329	-0.04	2.20E-07	-0.04	1.14E-06	-0.09	3.62E-06	-0.03	<b>1.98E-08</b>	-0.07	2.84E-07
CLDN5	chr22	19512942	cg05498726	-0.04	1.67E-05	-0.04	9.12E-05	-0.08	0.0005	-0.03	7.70E-06	-0.06	6.90E-05
CLDN5	chr22	19513006	cg11450827	-0.03	0.0004	-0.03	0.0024	-0.07	0.0008	-0.02	0.0003	-0.05	0.0003
CLDN5	chr22	19513008	cg17583256	-0.03	0.0002	-0.03	0.0056	-0.07	0.0011	-0.02	0.0022	-0.06	0.0004
CLDN5	chr22	19513017	cg16773741	-0.06	<b>5.06E-09</b>	-0.06	4.50E-05	-0.13	1.17E-05	-0.05	<b>2.77E-09</b>	-0.12	<b>1.81E-09</b>
CLDN5	chr22	19513078	cg14553765	-0.03	0.0010	-0.04	0.0010	-0.09	0.0006	-0.03	1.80E-05	-0.08	7.33E-06
CLDN5	chr22	19513176	cg00189989	-0.03	0.0004	-0.04	0.0005	-0.06	0.0041	-0.03	1.49E-05	-0.07	1.06E-05

Only CpG sites with a p-value < 0.001 for at least one cognitive domain are shown. Bonferroni threshold: 0.05/338,036=1.48E-07 (p-values below Bonferroni threshold in bold). EM: decline in episodic memory; PS: decline in perceptual speed; PO: decline in perceptual orientation; SM: decline in semantic memory; WM: decline in working memory. Adjusted for age at death, education, sex, ancestry, smoking status, post-mortem interval (PMI) and the first four principal components. Differential DNA methylation in the two lead signals (cg16773741 and cg05460329) was moderately correlated with cell type proportions (mainly PC2 (r=-0.32 for cg16773741 and r=-0.33 for cg05460329) and PC3 (r=-0.20 for cg16773741 and r=-0.14 for cg05460329), Table S21).

as implemented in the R package cate (35)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

## KEY RESOURCES TABLE

Resource Type	Specific Reagent or Resource	Source or Reference	Identifiers	Additional Information
Add additional rows as needed for each resource type	Include species and sex when applicable.	Include name of manufacturer, company, repository, individual, or research lab. Include PMID or DOI for references; use "this paper" if new.	Include catalog numbers, stock numbers, database IDs or accession numbers, and/or RRIDs. RRIDs are highly encouraged; search for RRIDs at <a href="https://scicrunch.org/resources">https://scicrunch.org/resources</a> .	Include any additional information or notes if necessary.
Deposited Data; Public Database	The methylation data was generated on prefrontal cortex samples from participants of the ROS/MAP studies using the Illumina HumanMethylation450 BeadChip. The brief description of the study and available files are located at the DOI listed. The individual IDAT files are located under Synapse ID "syn7357283" and the link provided in the "Identifiers" column of this spreadsheet.	doi: <a href="https://doi.org/10.7303/syn3157275">10.7303/syn3157275</a>	Individual IDAT files available at <a href="https://www.synapse.org/#!/Synapse:syn7357283">https://www.synapse.org/#!/Synapse:syn7357283</a>	
Deposited Data; Public Database	Rush Alzheimer's Disease Center Research Resource Sharing Hub	<a href="https://www.radc.rush.edu">https://www.radc.rush.edu</a>		Resource used to request phenotypic data.
Deposited Data; Public Database	MRC London Brain Bank for Neurodegenerative Disease	<a href="https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/mrc-london-neurodegenerative-diseases-brain-bank/">https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/mrc-london-neurodegenerative-diseases-brain-bank/</a>		
Deposited Data; Public Database	Ensembl gene predictions	<a href="http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/">http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/</a> .	ensGene, accessed on Apr-06-2014	
Software; Algorithm	The GAMuT software is available as an R script at the listed github URL.	<a href="https://github.com/epstein-software/GAMuT">https://github.com/epstein-software/GAMuT</a>	e0da206	